

SEQUENCE LISTING

<110> Merck & Co., Inc.
Aurismicchio, Luigi
La Monica, Nicola
Giannetti, Patrizia
Ciliberto, Gennaro

<120> SYNTHETIC GENE ENCODING RHESUS MONKEY
CARCINOEMBRYONIC ANTIGEN AND USES THEREOF

<130> ITR0058-PCT

<150> 60/497,201

<151> 2003-08-22

<160> 3

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 2118

<212> DNA

<213> Artificial Sequence

<220>

<223> codon-optimized rhesus monkey CEA

<400> 1

atgggcagcc	ccagcgcccc	cctgcaccgc	tggtgcatcc	cctggcagac	cctgctgctg	60
accgccagcc	tgctgacctt	ctggaacccc	cccaccaccg	cccagctgac	catcgagagc	120
cgccccttca	acgtggccga	gggcaaggag	gtgctgctgc	tggcccacaa	cgtgagccag	180
aacctgttcg	gctacatctg	gtacaagggc	gagcgcgtgg	acgccagccg	ccgcatcgcc	240
agctgcgtga	tccgcaccca	gcagatcacc	cccggccccg	cccacagcgg	ccgcgagacc	300
atcgacttca	acgccagcct	gctgatccac	aacgtgaccc	agagcgacac	cggcagctac	360
accatccagg	tgatcaagga	ggacctggtg	aacgaggagg	ccaccggcca	gttccgcgtg	420
taccocgagc	tgcccaagcc	ctacatcagc	agcaacaaca	gcaaccccg	ggaggacaag	480
gacgcgctgg	ccctgacctg	cgagcccag	accaggaca	ccacctacct	gtggtgggtg	540
aacaaccaga	gcctgcccgt	gagcccccg	ctggagctga	gcagcgacaa	ccgcaccctg	600
accgtgttca	acatcccccg	caacgacacc	accagctaca	agtgcgagac	ccagaacccc	660
gtgagcgtgc	gccgcagcga	ccccgtgacc	ctgaacgtgc	tgtacggccc	cgacgcccc	720
accatcagcc	ccctgaacac	cccctaccgc	gccggcgaga	acctgaacct	gacctgccac	780
gccgccagca	accccaccgc	ccagtacttc	tggttcgtga	acggcacctt	ccagcagagc	840
acccaggagc	tgttcatccc	caacatcacc	gtgaacaaca	gcggcagcta	catgtgccag	900
gcccacaaca	gcgccaccgg	cctgaaccgc	accaccgtga	ccgccatcac	cgtgtacgcc	960
gagctgcccc	agccctacat	caccagcaac	aacagcaacc	ccatcgagga	caaggacgcc	1020
gtgaccctga	cctgcgagcc	cgagaccag	gacaccacct	acctgtggtg	ggtgaacaac	1080
cagagcctga	gcgtgagcag	ccgcctggag	ctgagcaacg	acaaccgcac	cctgaccgtg	1140
ttcaacatcc	cccgcacaag	caccaccttc	tacgagtgcg	agaccagaa	ccccgtgagc	1200
gtgcgcgcga	gcgaccccg	gacctgaac	gtgctgtacg	gccccgacgc	ccccaccatc	1260
agccccctga	acacccccct	ccgcgcgggc	gagaacctga	acctgagctg	ccacgcggcc	1320
agcaaccocg	ccgcccagta	cagctggttc	gtgaacggca	ccttccagca	gagcaccacg	1380
gagctgttca	tccccaacat	caccgtgaac	aacagcggca	gctacatgtg	ccaggccccc	1440
aacagcgcca	ccggcctgaa	ccgcaccacc	gtgaccgcca	tcaccgtgta	cgtggagctg	1500
cccaagccct	acatcagcag	caacaacagc	aaccccatcg	aggacaagga	cgccgtgacc	1560
ctgacctgcg	agcccgctgg	cgagaacacc	acctacctgt	gggtgggtgaa	caaccagagc	1620
ctgagcgtga	gcccccgctt	gcagctgagc	aacggcaacc	gcacccctgac	cctgctgagc	1680
gtgaccgcga	acgacaccgg	cccctacgag	tcgggcatcc	agaacagcga	gagcgccaag	1740
cgcagcgacc	ccgtgacctt	gaacgtgacc	tacggccccg	acacccccat	catcagcccc	1800
cccgaacctga	gctaccgcag	cggcgccaac	ctgaacctga	gctgccacag	cgacagcaac	1860
cccagccccc	agtacagctg	gctgatcaac	ggcacctgc	gccagcacac	ccaggtgctg	1920
ttcatcagca	agatcaccag	caacaacagc	ggcgccctacg	cctgcttcgt	gagcaacctg	1980

```

gccaccggcc gcaacaacag catcgtgaag aacatcagcg tgagcagcgg cgacagcgcc 2040
cccgccagca gcggcctgag cgcccgcgcc accgtgggca tcatcatcgg catgctggtg 2100
ggcgtggccc tgatgtga                                     2118

```

<210> 2

<211> 705

<212> PRT

<213> Macaca Mulatta

<400> 2

```

Met Gly Ser Pro Ser Ala Pro Leu His Arg Trp Cys Ile Pro Trp Gln
 1      5      10      15
Thr Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
 20      25      30
Thr Ala Gln Leu Thr Ile Glu Ser Arg Pro Phe Asn Val Ala Glu Gly
 35      40      45
Lys Glu Val Leu Leu Leu Ala His Asn Val Ser Gln Asn Leu Phe Gly
 50      55      60
Tyr Ile Trp Tyr Lys Gly Glu Arg Val Asp Ala Ser Arg Arg Ile Gly
 65      70      75      80
Ser Cys Val Ile Arg Thr Gln Gln Ile Thr Pro Gly Pro Ala His Ser
 85      90      95
Gly Arg Glu Thr Ile Asp Phe Asn Ala Ser Leu Leu Ile His Asn Val
100      105      110
Thr Gln Ser Asp Thr Gly Ser Tyr Thr Ile Gln Val Ile Lys Glu Asp
115      120      125
Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
130      135      140
Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro Val Glu Asp Lys
145      150      155      160
Asp Ala Val Ala Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr Thr Tyr
165      170      175
Leu Trp Trp Val Asn Asn Gln Ser Leu Pro Val Ser Pro Arg Leu Glu
180      185      190
Leu Ser Ser Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro Arg Asn
195      200      205
Asp Thr Thr Ser Tyr Lys Cys Glu Thr Gln Asn Pro Val Ser Val Arg
210      215      220
Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp Ala Pro
225      230      235      240
Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn Leu Asn
245      250      255
Leu Thr Cys His Ala Ala Ser Asn Pro Thr Ala Gln Tyr Phe Trp Phe
260      265      270
Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile Pro Asn
275      280      285
Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His Asn Ser
290      295      300
Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val Tyr Ala
305      310      315      320
Glu Leu Pro Lys Pro Tyr Ile Thr Ser Asn Asn Ser Asn Pro Ile Glu
325      330      335
Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Glu Thr Gln Asp Thr
340      345      350
Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser Ser Arg
355      360      365
Leu Glu Leu Ser Asn Asp Asn Arg Thr Leu Thr Val Phe Asn Ile Pro
370      375      380
Arg Asn Asp Thr Thr Phe Tyr Glu Cys Glu Thr Gln Asn Pro Val Ser
385      390      395      400
Val Arg Arg Ser Asp Pro Val Thr Leu Asn Val Leu Tyr Gly Pro Asp
405      410      415
Ala Pro Thr Ile Ser Pro Leu Asn Thr Pro Tyr Arg Ala Gly Glu Asn
420      425      430

```

Leu Asn Leu Ser Cys His Ala Ala Ser Asn Pro Ala Ala Gln Tyr Ser
 435 440 445
 Trp Phe Val Asn Gly Thr Phe Gln Gln Ser Thr Gln Glu Leu Phe Ile
 450 455 460
 Pro Asn Ile Thr Val Asn Asn Ser Gly Ser Tyr Met Cys Gln Ala His
 465 470 475 480
 Asn Ser Ala Thr Gly Leu Asn Arg Thr Thr Val Thr Ala Ile Thr Val
 485 490 495
 Tyr Val Glu Leu Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro
 500 505 510
 Ile Glu Asp Lys Asp Ala Val Thr Leu Thr Cys Glu Pro Val Ala Glu
 515 520 525
 Asn Thr Thr Tyr Leu Trp Trp Val Asn Asn Gln Ser Leu Ser Val Ser
 530 535 540
 Pro Arg Leu Gln Leu Ser Asn Gly Asn Arg Ile Leu Thr Leu Leu Ser
 545 550 555 560
 Val Thr Arg Asn Asp Thr Gly Pro Tyr Glu Cys Gly Ile Gln Asn Ser
 565 570 575
 Glu Ser Ala Lys Arg Ser Asp Pro Val Thr Leu Asn Val Thr Tyr Gly
 580 585 590
 Pro Asp Thr Pro Ile Ile Ser Pro Pro Asp Leu Ser Tyr Arg Ser Gly
 595 600 605
 Ala Asn Leu Asn Leu Ser Cys His Ser Asp Ser Asn Pro Ser Pro Gln
 610 615 620
 Tyr Ser Trp Leu Ile Asn Gly Thr Leu Arg Gln His Thr Gln Val Leu
 625 630 635 640
 Phe Ile Ser Lys Ile Thr Ser Asn Asn Ser Gly Ala Tyr Ala Cys Phe
 645 650 655
 Val Ser Asn Leu Ala Thr Gly Arg Asn Asn Ser Ile Val Lys Asn Ile
 660 665 670
 Ser Val Ser Ser Gly Asp Ser Ala Pro Gly Ser Ser Gly Leu Ser Ala
 675 680 685
 Arg Ala Thr Val Gly Ile Ile Gly Met Leu Val Gly Val Ala Leu
 690 695 700
 Met
 705

<210> 3
 <211> 705
 <212> PRT
 <213> Macaca Mulatta

<400> 3
 Met Gly Ser Pro Ser Ala Pro Leu His Arg Trp Cys Ile Pro Trp Gln
 1 5 10 15
 Thr Leu Leu Leu Thr Ala Ser Leu Leu Thr Phe Trp Asn Pro Pro Thr
 20 25 30
 Thr Ala Gln Leu Thr Ile Glu Ser Arg Pro Phe Asn Val Ala Glu Gly
 35 40 45
 Lys Glu Val Leu Leu Leu Ala His Asn Val Ser Gln Asn Leu Phe Gly
 50 55 60
 Tyr Ile Trp Tyr Lys Gly Glu Arg Val Asp Ala Ser Arg Arg Ile Gly
 65 70 75 80
 Ser Cys Val Ile Arg Thr Gln Gln Ile Thr Pro Gly Pro Ala His Ser
 85 90 95
 Gly Arg Glu Thr Ile Asp Phe Asn Ala Ser Leu Leu Ile His Asn Val
 100 105 110
 Thr Gln Ser Asp Thr Gly Ser Tyr Thr Ile Gln Val Ile Lys Glu Asp
 115 120 125
 Leu Val Asn Glu Glu Ala Thr Gly Gln Phe Arg Val Tyr Pro Glu Leu
 130 135 140
 Pro Lys Pro Tyr Ile Ser Ser Asn Asn Ser Asn Pro Val Glu Asp Lys
 145 150 155 160

Asp	Ala	Val	Ala	Leu	Thr	Cys	Glu	Pro	Glu	Thr	Gln	Asp	Thr	Thr	Tyr
				165					170					175	
Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Pro	Val	Ser	Pro	Arg	Leu	Glu
			180					185					190		
Leu	Ser	Ser	Asp	Asn	Arg	Thr	Leu	Thr	Val	Phe	Asn	Ile	Pro	Arg	Asn
		195					200					205			
Asp	Thr	Thr	Ser	Tyr	Lys	Cys	Glu	Thr	Gln	Asn	Pro	Val	Ser	Val	Arg
	210				215						220				
Arg	Ser	Asp	Pro	Val	Thr	Leu	Asn	Val	Leu	Tyr	Gly	Pro	Asp	Ala	Pro
	225				230					235					240
Thr	Ile	Ser	Pro	Leu	Asn	Thr	Pro	Tyr	Arg	Ala	Gly	Glu	Asn	Leu	Asn
				245					250					255	
Leu	Thr	Cys	His	Ala	Ala	Ser	Asn	Pro	Thr	Ala	Gln	Tyr	Phe	Trp	Phe
			260					265					270		
Val	Asn	Gly	Thr	Phe	Gln	Gln	Ser	Thr	Gln	Glu	Leu	Phe	Ile	Pro	Asn
		275					280					285			
Ile	Thr	Val	Asn	Asn	Ser	Gly	Ser	Tyr	Met	Cys	Gln	Ala	His	Asn	Ser
	290					295					300				
Ala	Thr	Gly	Leu	Asn	Arg	Thr	Thr	Val	Thr	Ala	Ile	Thr	Val	Tyr	Ala
	305				310					315					320
Glu	Leu	Pro	Lys	Pro	Tyr	Ile	Thr	Ser	Asn	Asn	Ser	Asn	Pro	Ile	Glu
				325					330					335	
Asp	Lys	Asp	Ala	Val	Thr	Leu	Thr	Cys	Glu	Pro	Glu	Thr	Gln	Asp	Thr
			340					345					350		
Thr	Tyr	Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Ser	Val	Ser	Ser	Arg
		355					360					365			
Leu	Glu	Leu	Ser	Asn	Asp	Asn	Arg	Thr	Leu	Thr	Val	Phe	Asn	Ile	Pro
	370					375					380				
Arg	Asn	Asp	Thr	Thr	Phe	Tyr	Glu	Cys	Glu	Thr	Gln	Asn	Pro	Val	Ser
	385				390					395					400
Val	Arg	Arg	Ser	Asp	Pro	Val	Thr	Leu	Asn	Val	Leu	Tyr	Gly	Pro	Asp
				405					410					415	
Ala	Pro	Thr	Ile	Ser	Pro	Leu	Asn	Thr	Pro	Tyr	Arg	Ala	Gly	Glu	Asn
			420					425					430		
Leu	Asn	Leu	Ser	Cys	His	Ala	Ala	Ser	Asn	Pro	Ala	Ala	Gln	Tyr	Phe
		435				440						445			
Trp	Phe	Val	Asn	Gly	Thr	Phe	Gln	Gln	Ser	Thr	Gln	Glu	Leu	Phe	Ile
	450					455					460				
Pro	Asn	Ile	Thr	Val	Asn	Asn	Ser	Gly	Ser	Tyr	Met	Cys	Gln	Ala	His
	465				470					475					480
Asn	Ser	Ala	Thr	Gly	Leu	Asn	Arg	Thr	Thr	Val	Thr	Ala	Ile	Thr	Val
				485					490					495	
Tyr	Val	Glu	Leu	Pro	Lys	Pro	Tyr	Ile	Ser	Ser	Asn	Asn	Ser	Asn	Pro
		500						505					510		
Ile	Glu	Asp	Lys	Asp	Ala	Val	Thr	Leu	Thr	Cys	Glu	Pro	Val	Ala	Glu
		515					520					525			
Asn	Thr	Thr	Tyr	Leu	Trp	Trp	Val	Asn	Asn	Gln	Ser	Leu	Ser	Val	Ser
	530				535						540				
Pro	Arg	Leu	Gln	Leu	Ser	Asn	Gly	Asn	Arg	Ile	Leu	Thr	Leu	Leu	Ser
	545				550					555					560
Val	Thr	Arg	Asn	Asp	Thr	Gly	Pro	Tyr	Glu	Cys	Gly	Ile	Gln	Asn	Ser
				565					570					575	
Glu	Ser	Ala	Lys	Arg	Ser	Asp	Pro	Val	Thr	Leu	Asn	Val	Thr	Tyr	Gly
			580					585					590		
Pro	Asp	Thr	Pro	Ile	Ile	Ser	Pro	Pro	Asp	Leu	Ser	Tyr	Arg	Ser	Gly
		595					600					605			
Ala	Asn	Leu	Asn	Leu	Ser	Cys	His	Ser	Asp	Ser	Asn	Pro	Ser	Pro	Gln
		610				615					620				
Tyr	Ser	Trp	Leu	Ile	Asn	Gly	Thr	Leu	Arg	Gln	His	Thr	Gln	Val	Leu
	625				630					635					640
Phe	Ile	Ser	Lys	Ile	Thr	Ser	Asn	Asn	Asn	Gly	Ala	Tyr	Ala	Cys	Phe
				645					650					655	
Val	Ser	Asn	Leu	Ala	Thr	Gly	Arg	Asn	Asn	Ser	Ile	Val	Lys	Asn	Ile
			660					665					670		

Ser	Val	Ser	Ser	Gly	Asp	Ser	Ala	Pro	Gly	Ser	Ser	Gly	Leu	Ser	Ala
		675					680					685			
Arg	Ala	Thr	Val	Gly	Ile	Ile	Ile	Gly	Met	Leu	Val	Gly	Val	Ala	Leu
	690					695					700				
Met															
705															